

J. Higgins

1641

PAGE: 1

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/961,083

DATE: 01/29/1999
TIME: 17:50:39

INPUT SET: S30408.raw

This Raw Listing contains the General
Information Section and up to the first 5 pages.

SEQUENCE LISTING

(1) General Information:

(i) APPLICANT: Choi et. al.

(ii) TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines

(iii) NUMBER OF SEQUENCES: 452

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Human Genome Sciences, Inc.

(B) STREET: 9410 Key West Avenue

(C) CITY: Rockville

(D) STATE: Maryland

(E) COUNTRY: USA

(F) ZIP: 20850

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

(B) COMPUTER: HP Vectra 486/33

(C) OPERATING SYSTEM: MSDOS version 6.2

(D) SOFTWARE: ASCII Text

ENTERED

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/961,083DATE: 01/29/1999
TIME: 17:50:40

INPUT SET: S30408.raw

47 (vi) CURRENT APPLICATION DATA:
48
49 (A) APPLICATION NUMBER:
50
51 (B) FILING DATE:
52
53 (C) CLASSIFICATION:
54
55
56
57 (vii) PRIOR APPLICATION DATA:
58
59 (A) APPLICATION NUMBER:
60
61 (B) FILING DATE:
62
63
64
65 (viii) ATTORNEY/AGENT INFORMATION:
66
67 (A) NAME: Brookes, A. Anders
68
69 (B) REGISTRATION NUMBER: 36,373
70
71 (C) REFERENCE/DOCKET NUMBER: PB340P2
72
73
74
75 (vi) TELECOMMUNICATION INFORMATION:
76
77 (A) TELEPHONE: (301) 309-8504
78
79 (B) TELEFAX: (301) 309-8512
80
81
82
83
84
85
86

87 (2) INFORMATION FOR SEQ ID NO: 1:
88
89 (i) SEQUENCE CHARACTERISTICS:
90 (A) LENGTH: 1999 base pairs
91 (B) TYPE: nucleic acid
92 (C) STRANDEDNESS: double
93 (D) TOPOLOGY: linear
94
95
96
97 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
98
99 TAAAATCTAC GACAATAAAA ATCAACTCAT TGCTGACTTG GGTTCCTGAAC GCCGCGTCAA

RAW SEQUENCE LISTING
PATENT APPLICATION *US/08/961,083*

DATE: 01/29/1999
TIME: 17:50:40

INPUT SET: S30408.raw

100							
101	TGCCCAAGCT	AATGATATTC	CCACAGATTT	GGTTAAGGCA	ATCGTTTCTA	TCGAAGACCA	120
102							
103	TCGCTTCTTC	GACCACAGGG	GGATTGATAC	CATCCGTATC	CTGGGAGCTT	TCTTGCGCAA	180
104							
105	TCTGCAAAGC	AATTCCCTCC	AAGGTGGATC	AACTCTCACC	CAACAGTTGA	TTAAGTTGAC	240
106							
107	TTACTTTTCA	ACTTCGACTT	CCGACCAGAC	TATTTCTCGT	AAGGCTCAGG	AAGCTTGGTT	300
108							
109	AGCGATTGAG	TTAGAACAAA	AAGCAACCAA	GCAAGAAATC	TTGACCTACT	ATATAAATAA	360
110							
111	GGTCTACATG	TCTAATGGGA	ACTATGGAAT	GCAGACAGCA	GCTCAAAACT	ACTATGGTAA	420
112							
113	AGACCTCAAT	AATTTAAGTT	TACCTCAGTT	AGCCTTGCTG	GCTGGAATGC	CTCAGGCACC	480
114							
115	AAACCAATAT	GACCCCTATT	CACATCCAGA	AGCAGCCCAA	GACCGCCGAA	ACTTGGTCTT	540
116							
117	ATCTGAAATG	AAAAATCAAG	GCTACATCTC	TGCTGAACAG	TATGAGAAAG	CAGTCAATAC	600
118							
119	ACCAATTACT	GATGGACTAC	AAAGTCTCAA	ATCAGCAAGT	AATTACCCTG	CTTACATGGA	660
120							
121	TAATTACCTC	AAGGAAGTCA	TCAATCAAGT	TGAAGAAGAA	ACAGGCTATA	ACCTACTCAC	720
122							
123	AACTGGGATG	GATGTCTACA	CAAATGTAGA	CCAAGAAGCT	CAAAAACATC	TGTGGGATAT	780
124							
125	TTACAATACA	GACGAATACG	TTGCCTATCC	AGACGATGAA	TTGCAAGTCG	CTTCTACCAT	840
126							
127	TGTTGATGTT	TCTAACGGTA	AAGTCATTGC	CCAGCTAGGA	GCACGCCATC	AGTCAAGTAA	900
128							
129	TGTTTCCTTC	GGAATTAACC	AAGCAGTAGA	AACAAACCGC	GACTGGGGAT	CAACTATGAA	960
130							
131	ACCGATCACA	GACTATGCTC	CTGCCTTGGA	GTACGGTGTC	TACGATTCAA	CTGCTACTAT	1020
132							
133	CGTTCACGAT	GAGCCCTATA	ACTACCCCTGG	GACAAATACT	CCTGTTTATA	ACTGGGATAG	1080
134							
135	GGGCTACTTT	GGCAACATCA	CCTTGCAATA	CGCCCTGCAA	CAATCGCGAA	ACGTCCCAGC	1140
136							
137	CGTGGAAACT	CTAAACAAGG	TCGGACTCAA	CCGCGCCAAG	ACTTTCCTAA	ATGGTCTAGG	1200
138							
139	AATCGACTAC	CCAAGTATTC	ACTACTCAAA	TGCCATTTCA	AGTAACACAA	CCGAATCAGA	1260
140							
141	CAAAAAATAT	GGAGCAAGTA	GTGAAAAGAT	GGCTGCTGCT	TACGCTGCCT	TTGCAAATGG	1320
142							
143	TGGAAC TTAC	TATAAACC AA	TGTATATCCA	TAAAGTCGTC	TTTAGTGATG	GGAGTGAAAA	1380
144							
145	AGAGTTCTCT	AATGTCGGAA	CTCGTGCCAT	GAAGGAAACG	ACAGCCTATA	TGATGACCGA	1440
146							
147	CATGATGAAA	ACAGTCTTGA	CTTATGGAAC	TGGACGAAAT	GCCTATCTTG	CTTGGCTCCC	1500
148							
149	TCAGGCTGGT	AAAACAGGAA	CCTCTAACTA	TACAGACGAG	GAAATTGAAA	ACCACATCAA	1560
150							
151	GACCTCTCAA	TTTGTAGCAC	CTGATGAACT	ATTTGCTGGC	TATACGCGTA	AATATTCAAT	1620
152							

RAW SEQUENCE LISTING PATENT APPLICATION US/08/961,083

DATE: 01/29/1999
TIME: 17:50:40

INPUT SET: S30408.raw

```

153 GGCTGTATGG ACAGGCTATT CTAACCGTCT GACACCACTT GTAGGCAATG GCCTTACGGT 1680
154
155 CGCTGCCAAA GTTTACCGCT CTATGATGAC CTACCTGTCT GAAGGAAGCA ATCCAGAAGA 1740
156
157 TTGGAATATA CCAGAGGGGC TCTACAGAAA TGGAGAATTC GTATTTAAAA ATGGTGCTCG 1800
158
159 TTCTACGTGG AACTCACCTG CTCCACAACA ACCCCCATCA ACTGAAAGTT CAAGCTCATC 1860
160
161 ATCAGATAGT TCAACTTCAC AGTCTAGCTC AACCCTCCA AGCACAAATA ATAGTACGAC 1920
162
163 TACCAATCCT AACAATAATA CGCAACAATC AAATACAACC CCTGATCAAC AAAATCAGAA 1980
164
165 TCCTCAACCA GCACAACCA 1999
166

```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 666 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

182 Lys Ile Tyr Asp Asn Lys Asn Gln Leu Ile Ala Asp Leu Gly Ser Glu
183 1 5 10 15
184
185 Arg Arg Val Asn Ala Gln Ala Asn Asp Ile Pro Thr Asp Leu Val Lys
186 20 25 30
187
188 Ala Ile Val Ser Ile Glu Asp His Arg Phe Phe Asp His Arg Gly Ile
189 35 40 45
190
191 Asp Thr Ile Arg Ile Leu Gly Ala Phe Leu Arg Asn Leu Gln Ser Asn
192 50 55 60
193
194 Ser Leu Gln Gly Gly Ser Thr Leu Thr Gln Gln Leu Ile Lys Leu Thr
195 65 70 75 80
196
197 Tyr Phe Ser Thr Ser Thr Ser Asp Gln Thr Ile Ser Arg Lys Ala Gln
198 85 90 95
199
200 Glu Ala Trp Leu Ala Ile Gln Leu Glu Gln Lys Ala Thr Lys Gln Glu
201 100 105 110
202
203 Ile Leu Thr Tyr Tyr Ile Asn Lys Val Tyr Met Ser Asn Gly Asn Tyr
204 115 120 125
205

```

RAW SEQUENCE LISTING PATENT APPLICATION US/08/961,083

DATE: 01/29/1999
TIME: 17:50:41

INPUT SET: S30408.raw

206	Gly Met Gln Thr Ala Ala Gln Asn Tyr Tyr Gly Lys Asp Leu Asn Asn
207	130 135 140
208	
209	Leu Ser Leu Pro Gln Leu Ala Leu Leu Ala Gly Met Pro Gln Ala Pro
210	145 150 155 160
211	
212	Asn Gln Tyr Asp Pro Tyr Ser His Pro Glu Ala Ala Gln Asp Arg Arg
213	165 170 175
214	
215	Asn Leu Val Leu Ser Glu Met Lys Asn Gln Gly Tyr Ile Ser Ala Glu
216	180 185 190
217	
218	Gln Tyr Glu Lys Ala Val Asn Thr Pro Ile Thr Asp Gly Leu Gln Ser
219	195 200 205
220	
221	Leu Lys Ser Ala Ser Asn Tyr Pro Ala Tyr Met Asp Asn Tyr Leu Lys
222	210 215 220
223	
224	Glu Val Ile Asn Gln Val Glu Glu Glu Thr Gly Tyr Asn Leu Leu Thr
225	225 230 235 240
226	
227	Thr Gly Met Asp Val Tyr Thr Asn Val Asp Gln Glu Ala Gln Lys His
228	245 250 255
229	
230	Leu Trp Asp Ile Tyr Asn Thr Asp Glu Tyr Val Ala Tyr Pro Asp Asp
231	260 265 270
232	
233	Glu Leu Gln Val Ala Ser Thr Ile Val Asp Val Ser Asn Gly Lys Val
234	275 280 285
235	
236	Ile Ala Gln Leu Gly Ala Arg His Gln Ser Ser Asn Val Ser Phe Gly
237	290 295 300
238	
239	Ile Asn Gln Ala Val Glu Thr Asn Arg Asp Trp Gly Ser Thr Met Lys
240	305 310 315 320
241	
242	Pro Ile Thr Asp Tyr Ala Pro Ala Leu Glu Tyr Gly Val Tyr Asp Ser
243	325 330 335
244	
245	Thr Ala Thr Ile Val His Asp Glu Pro Tyr Asn Tyr Pro Gly Thr Asn
246	340 345 350
247	
248	Thr Pro Val Tyr Asn Trp Asp Arg Gly Tyr Phe Gly Asn Ile Thr Leu
249	355 360 365
250	
251	Gln Tyr Ala Leu Gln Gln Ser Arg Asn Val Pro Ala Val Glu Thr Leu
252	370 375 380
253	
254	Asn Lys Val Gly Leu Asn Arg Ala Lys Thr Phe Leu Asn Gly Leu Gly
255	385 390 395 400
256	
257	Ile Asp Tyr Pro Ser Ile His Tyr Ser Asn Ala Ile Ser Ser Asn Thr
258	405 410 415

PAGE: 1

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION US/08/961,083

DATE: 01/29/1999
TIME: 17:50:41

INPUT SET: S30408.raw

Line

Error

Original Text